Getting Started with R

Bi Zhao

Department of Political Science
Purdue University

zhao76@purdue.edu

August 17, 2019

Overview

Before we start, go to this Dropbox folder and download all the materials.

Getting Started

- To launch RStudio, locate the icon that is in the Start menu (Windows) or Applications folder (Mac).
- Users will note that there are a few basic pull-down functions and menus available.
- However, the great bulk of operations within R are completed using user-entered code.

RStudio IDE

- R script
- Console for the R session
- Files, Packages, Plots
- Environment, History

Setting working directory

- The commands you need:
  getwd()
  setwd()
- Let's go to RStudio and set your working directory

Rules for Entering Code

- Commands are case sensitive: For example, the command var and VAR have two different meanings.
- R treats anything following the pound sign #as a comment. This is useful for making notes about commands, etc.
- An object name needs to start with an alphabetic character, but may contain numeric characters afterwards. Example: x., monkey1
- Don't name your objects things like "mean"or "sum" or "7" since those are things that R already has pre-packaged.
- You can use the up/down arrow keys to scroll through preceding/subsequent commands.

Arithmetic commands

- Basic arithmetic operations are coded in R: for example, addition (+), subtraction (-), multiplication (*), division (/).
- Parentheses denote order of operations. Example: $(3 + 5/78)^3 * 7$

Data types

- R supports a wide variety of data types. You can use class() command to find out the type.
- Scalars
    - Number
    - Logical value
    - Character (string)
- Vector
- Matrix
- Data frame
- List

Objects

- Objects: all data types, operations, functions
- Defined by <-

Load data into R

- R can read different types of data.
- Before you load the data, make sure you have the correct working directory.

R packages

- R packages are collections of functions, sample data, and compiled codes.
- They help you to execute certain functions.
- For first time use, install the package.
- Once installed, you need to load the package each time you are in a new session.

Basic R Functions

- R has many programmed functions that you can use to explore your data:
  sum, mean, median, range, max, min, sd, var, etc.

Getting a feel of data in R

- Let's play with the datasets we just downloaded from Dropbox.

Sweden data

- country: Abbreviated country name
- period: Period during which data are collected
- age: Age group
- births: Number of births in thousands
- deaths: Number of deaths in thousands
- py.men: Men in thousands
- py.women: Women in thousands

- year: Election year
- VEP: Voting eligible population in thousands
- VAP: Voting age population in thousands
- total: Total ballots cast for highest office in thousands

SWIRL

- Swirl is an interactive learning package in R.
- First, install the package;
- library(swirl)
- swirl()

Questions?